

Neuropsychological Rehabilitation

An International Journal

ISSN: 0960-2011 (Print) 1464-0694 (Online) Journal homepage: <https://www.tandfonline.com/loi/pnrh20>

Comparison of methods for estimating premorbid intelligence

Peter Bright & Ian van der Linde

To cite this article: Peter Bright & Ian van der Linde (2020) Comparison of methods for estimating premorbid intelligence, *Neuropsychological Rehabilitation*, 30:1, 1-14, DOI: [10.1080/09602011.2018.1445650](https://doi.org/10.1080/09602011.2018.1445650)

To link to this article: <https://doi.org/10.1080/09602011.2018.1445650>



© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 12 Mar 2018.



[Submit your article to this journal](#)



Article views: 20129



[View related articles](#)



[View Crossmark data](#)



[Citing articles: 30 View citing articles](#)



OPEN ACCESS



Comparison of methods for estimating premorbid intelligence

Peter Bright^{a,b} and Ian van der Linde^{b,c}

^aDepartment of Psychology, Anglia Ruskin University, Cambridge, UK; ^bVision & Eye Research Unit (VERU), Postgraduate Medical Institute, Anglia Ruskin University, Cambridge, UK; ^cDepartment of Computing & Technology, Anglia Ruskin University, Cambridge, UK

ABSTRACT

To evaluate impact of neurological injury on cognitive performance it is typically necessary to derive a baseline (or “premorbid”) estimate of a patient’s general cognitive ability prior to the onset of impairment. In this paper, we consider a range of common methods for producing this estimate, including those based on current best performance, embedded “hold/no-hold” tests, demographic information, and word reading ability. Ninety-two neurologically healthy adult participants were assessed on the full Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV; Wechsler, D. (2008). *Wechsler Adult Intelligence Scale* (4th ed.). San Antonio, TX: Pearson Assessment.) and on two widely used word reading tests: National Adult Reading Test (NART; Nelson, H. E. (1982). *National Adult Reading Test (NART): For the assessment of premorbid intelligence in patients with dementia: Test manual*. Windsor: NFER-Nelson.; Nelson, H. E., & Willison, J. (1991). *National Adult Reading Test (NART)*. Windsor: NFER-Nelson.) and Wechsler Test of Adult Reading (WTAR; Wechsler, D. (2001). *Wechsler Test of Adult Reading: WTAR*. San Antonio, TX: Psychological Corporation.). Our findings indicate that reading tests provide the most reliable and precise estimates of WAIS-IV full-scale IQ, although the addition of demographic data provides modest improvement. Nevertheless, we observed considerable variability in correlations between NART/WTAR scores and individual WAIS-IV indices, which indicated particular usefulness in estimating more crystallised premorbid abilities (as represented by the verbal comprehension and general ability indices) relative to fluid abilities (working memory and perceptual reasoning indices). We discuss and encourage the development of new methods for improving premorbid estimates of cognitive abilities in neurological patients.

ARTICLE HISTORY Received 16 October 2017; Accepted 20 February 2018

KEYWORDS Neuropsychological assessment; premorbid intelligence; NART; WTAR; WAIS-IV

Introduction

Several approaches have been devised to estimate premorbid cognitive ability in neurological patients. These include *best performance* (Lezak, 1995), “*hold/no-hold*”

CONTACT Peter Bright  peter.bright@anglia.ac.uk  Department of Psychology, Anglia Ruskin University, East Road, Cambridge CB1 1PT, UK

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

(Wechsler, 1958), demographics (e.g., Barona, Reynolds, & Chastain, 1984; Crawford & Allan, 1997), reading ability (e.g., Nelson, 1982; Nelson & Willison, 1991; Wechsler, 2001), and combinations thereof (e.g., Crawford, Nelson, Blackmore, Cochrane, & Allan, 1990; Vanderploeg, Schinka, & Axelrod, 1996). The appropriateness of a given approach is likely to depend on the patient under investigation, but those based on reading ability/word knowledge are among the most widely employed, particularly in North America, UK and Australia (e.g., Crawford, Stewart, Cochrane, Parker, & Besson, 1989; Mathias, Bowden, & Barrett-Woodbridge, 2007; Skilbeck, Dean, Thomas, & Slatyer, 2013). However, there are few published methods currently available that have been standardised against the most recent revision of the Wechsler Adult Intelligence Scale (WAIS-IV; Wechsler, 2008). In this study, we compare the precision of a range of approaches for estimating WAIS-IV full-scale IQ (FSIQ) and constituent indices and offer new combined methods that clinicians and researchers may wish to consider adopting in their work.

A large body of evidence suggests that scores on tests requiring the reading of phonetically irregular words, such as the National Adult Reading Test (NART; Nelson, 1982; Nelson & Willison, 1991) and Wechsler Test of Adult Reading (WTAR; Wechsler, 2001), are highly correlated with measured intelligence in healthy populations (e.g., Bright, Jaldow, & Kopelman, 2002; Bright, Hale, Gooch, Myhill, & van der Linde, 2016; Crawford, Deary, Starr, & Whalley, 2001; Nelson & O'Connell, 1978), and that reading ability, particularly of irregular words, is resistant to neurological impairment and age-related cognitive decline (for reviews see Franzen, Burgess, & Smith-Seemiller, 1997; Lezak, Howieson, Bigler, & Tranel, 2012). Although the relative utility and accuracy of these tests for many neurological conditions is unknown, Bright et al. (2002) provided evidence that the use of the NART is justified in patients with frontal lobe damage, Korsakoff syndrome, and mild or moderate stages of Alzheimer's disease, and that this test outperforms demographic-derived estimates, with no additional benefit to be gained from a combination of the two methods. However, it is widely accepted that such tests are likely to provide the most reliable premorbid estimates in the average range, whilst overestimating IQ in those with very low scores and underestimating those with very high scores (see, for example, Bright et al., 2016; Nelson & Willison, 1991).

Although the NART and WTAR are among the most popular instruments for estimating premorbid WAIS IQ, only the former has been standardised against the most recent (fourth revision) of the WAIS battery (Bright et al., 2016). The Test of Premorbid Functioning (TOPF; Pearson, 2009; Wechsler, 2011), proposed as a replacement for the WTAR, has been standardised against WAIS-IV, but has not been widely adopted to date (at least for research purposes). Figure 1 provides an indication of comparative popularity of NART, WTAR and TOPF in research year-by-year. Although it is important to note that total citation counts will be biased towards longer established tests, they clearly demonstrate continued use of the NART and the WTAR, despite some indication that the TOPF is gaining popularity.

Best performance approaches to estimating premorbid ability are based upon the assumption that the tests in which patients accrue the highest score are likely to reflect relatively intact function, and therefore provide a baseline ability level against which current functioning can be compared. Typically, the clinician infers general premorbid ability on the basis of the one or two best WAIS-IV subtest scores, but given the considerable variability among the subtests observed in healthy populations, it is acknowledged that this approach is likely to significantly overestimate premorbid ability (Franzen et al., 1997; Griffin, Mindt, Rankin, Ritchie, & Scott, 2002; Mortensen,

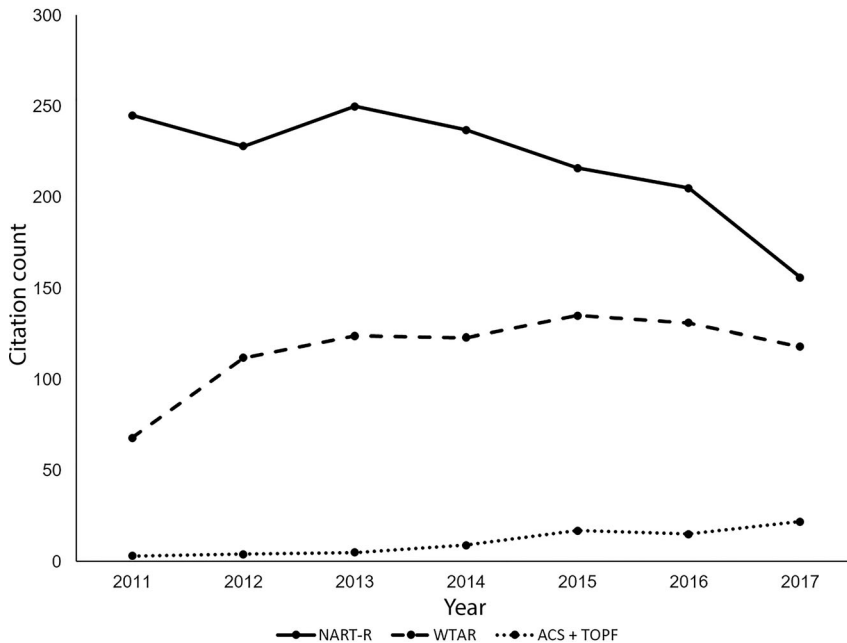


Figure 1. Number of academic publications in which NART-R (solid line), WTAR (dashed line) and Advanced Clinical Solutions/Test of Premorbid Functioning (ACS/TOPF) (dotted line) neuropsychological tests were cited for each year from 2011 to October 2017. Google Scholar (5 October 5 2017) citation counts based on [Nelson and Willison (1991). *National Adult Reading Test (NART)*. NFER-Nelson] for NART-R; [Wechsler (2001). *Wechsler Test of Adult Reading: WTAR*. Psychological Corporation] for WTAR, and combined counts from [Pearson (2009). *Advanced Clinical Solutions for WAIS-IV and WMS-IV: Administration and scoring manual*. The Psychological Corporation, San Antonio] and [Wechsler (2011). *Test of Premorbid Functioning. UK Version (TOPF UK)*. UK: Pearson Corporation] for ACS/TOPF.

Gade, & Reinisch, 1991; Reynolds, 1997). Some authors have, in response to this problem, developed a “correction” to be applied to such estimates that uses demographic (and other) information, but have not satisfactorily resolved the tendency towards premorbid IQ overestimation (Powell, Brossart, & Reynolds, 2003).

In the WAIS batteries, Vocabulary, Matrix Reasoning, Information and Picture Completion subtests are those least likely to be affected by brain damage (e.g., Donders, Tulsky, & Zhu, 2001; Wechsler, 1997), and are therefore considered to be embedded “hold” tests, against which those subtests more sensitive to damage (the “no-hold” tests) can be compared. Lezak (2012) suggests that Vocabulary and Information are the best/classic “hold subtests”. Using this approach, premorbid ability can be inferred on the basis of current WAIS performance – an advantage to the extent that like is compared with like. However, such WAIS subtests may be more sensitive to neurological damage than standalone tests of word reading/knowledge, such as the NART and WTAR (Franzen et al., 1997; Reynolds, 1997). Furthermore, the calculation of a premorbid IQ estimate on the basis of a subset of the same tests used to calculate current IQ suggests a psychometric flaw, in which there is very likely to be high predictive accuracy in healthy populations but questionable validity when applied in neurological patients. For example, Powell et al. (2003) provide evidence that the Oklahoma Premorbid Intelligence Estimate (OPIE; Scott, Krull, Williamson, Adams, & Iverson, 1997), based on

combined “hold” WAIS subtest and demographic information, produces estimates in cognitively impaired patients which may be closer to their current than premorbid IQ (i.e., the method underestimates patient deficit). Finally, the hold/no-hold approach, like best performance, requires that we accept the assumption that neurologically healthy populations perform similarly across all subtests. However, the weight of evidence is not consistent with this view.

In the present study, we examine the accuracy with which the NART and WTAR predict intelligence on the most recent revision of the Wechsler Adult Intelligence Scale (WAIS-IV), using a large sample of neurologically healthy participants ($n = 92$). We also consider an abbreviated form of the NART (mini-NART, McGrory, Austin, Shenkin, Starr, & Deary, 2015), developed in order to expedite the test and remove words that provide little additional predictive power. Furthermore, we assess whether a combination of NART/WTAR and demographic information improves predictive accuracy and compare NART/WTAR performance against the WAIS-IV embedded “hold” tests as measures of WAIS-IV FSIQ. Our overall aim was to establish which method, or combination of methods, offers the most accurate prediction of WAIS-IV FSIQ and its constituent indices.

Method

Participants

An opportunity sample of 100 neurologically healthy adults (mean age 40 years; range 18 to 70; SD 16.78) were recruited primarily from university campuses in Cambridge and London, local retail environments and via social media, of which eight participants failed to complete one or more tests and were excluded from all analyses. There were no missing data across the sample of 92 participants for any variable, with the exception of social class (missing for 14 participants, as indicated in Table 1). Table 1 provides demographic and WAIS-IV FSIQ data. All were British nationals, with English as the

Table 1. WAIS-IV performance and demographics.

	Mean	SD	<i>N</i>	Sample proportion (%)
Age	40.00	16.78	92	100
FSIQ	108.52	12.71	92	100
80–90			7	8
91–100			15	16
101–110			30	33
111–120			23	25
121+			17	18
Social class^a	3.15	1.41	78	85
I Professional			8	10
II Managerial/technical			15	19
III Skilled non-manual			26	33
IV Partly skilled			16	21
V Unskilled			13	17
Education	2.30	.98	92	100
I GCSE/equivalent or below			25	27
II A level/equivalent			23	25
III Undergraduate degree			35	38
IV Postgraduate degree			9	10

Note: Full sample statistics are indicated in bold. Occupation information unavailable for 14 participants; education data lists maximum qualification obtained (or in progress). Typically, school leaving age of 16 corresponds to level I, 18 to level II; levels III and IV included participants currently undertaking that level of study. WAIS-IV, Wechsler Adult Intelligence Scale – Fourth Edition.

first language, and with normal/corrected-to-normal vision and hearing. Participants self-declared that they had no history of neurological or psychiatric disorder. Extensive training in the administration and scoring of all tests was provided to three research assistants over several days by the lead author, and the testing sessions were closely monitored and supervised to ensure full compliance with the standardised administration and scoring procedures. All participants were recruited and tested between 2013 and 2016, in a UK university setting.

Materials and procedure

Demographic information was recorded (age, gender, years of education, occupation), with social class determined by occupation using the Office of Population, Censuses and Surveys (1980) British classification, which ranges from 1 (professional) to 5 (unskilled). The British NART, WTAR and WAIS-IV were then administered (in that order) according to standardised instructions. Data for the 23 items comprising the mini-NART (McGrory et al., 2015) were extracted to provide an overall score on this abbreviated version of the test. The WAIS-IV supplementary tests were administered to all participants at the end of the session but will not be reported here. Procedures were approved by the University ethics panel and followed the tenets of the Declaration of Helsinki. Data were collected from all participants in one session.

Results

Participant demographics and WAIS-IV performance are shown in Table 1. The FSIQ range was 80 to 150, with an arithmetic mean of 108.52 and standard deviation of 12.71. All levels of occupation and education were represented.

Best performance

To determine the viability of using a straightforward best performance approach to estimating premorbid IQ, we assessed variability in performance across WAIS-IV subtests and indices in our neurologically healthy sample. Four separate indices were introduced with WAIS-IV, replacing the verbal and performance subscales included in previous versions of the test battery: Verbal Comprehension (VCI), Perceptual Reasoning (PRI), Working Memory (WMI) and Processing speed (PSI). Additionally, scores on the VCI and PRI subtests contribute to a General Ability Index (GAI), typically employed in cases in which disproportionate working memory and/or processing speed difficulties complicate the interpretation of FSIQ (Wechsler, 2008).

Mean performance across the subtests was generally similar, with only four significant differences, following Bonferroni correction for multiple comparisons. Scaled scores were higher for Information in comparison with Digit Span ($p = .046$), Coding ($p = .041$) and Similarities ($p < .01$), and for Block Design in comparison to Similarities ($p = .038$). No differences were observed among the index scores ($p > .05$ in all cases). Despite the modest disparity among the subtest and index means, marked within-subject variability in performance was found. To illustrate this, we recorded the lowest and highest index scores for each participant. A comparison of these means in our sample revealed a 22.62 point discrepancy (mean lowest = 95.27; highest = 117.89). Similarly, a comparison of participants' mean lowest subtest scaled score

(7.85) against their highest subtest scaled score (14.77) revealed a mean difference of 6.92 scaled points. Such variability in neurologically healthy participants renders estimation of premorbid IQ using a straightforward best performance approach problematic, and likely to produce markedly inflated predicted scores.

Hold vs. no-hold

To address the viability of the hold vs. no-hold approach to estimating premorbid cognitive ability, we selected “hold” and “no-hold” subtests according to Lezak’s (2012) categorisation. Typically, Vocabulary and Information are employed as hold tests because they are considered disproportionately resistant to neurological and psychological impairment (e.g., Groth-Marnat & Wright, 2016; Lezak et al., 2012). Less commonly, Picture Completion (now a supplementary rather than core test) and Matrix Reasoning are also employed but will not be included here. By extension, the remaining core subtests measure “no-hold” abilities (i.e., those most susceptible to neurocognitive impairment), but the most commonly used are Block Design, Digit Span, Arithmetic and/or Coding (Groth-Marnat & Wright, 2016; Wechsler, 1958). Anecdotally, and in clinical practice, two tests are commonly selected to provide a comparator against hold performance (Block Design and Digit Span). Table 2 presents linear correlations between hold and no-hold tests, along with combined measures. Paired *t*-tests (two-tailed) revealed significant differences between hold and no-hold combined measurements. Correlation coefficients, although significant, were relatively small, even though statistical power ($1 - \beta$) in all cases exceeded .8 (two-tailed). For example, the shared variance (r^2) between Vocabulary and Block Design scaled scores was less than 10%, rising to 12% for the combined hold measure. Correlations between the combined hold and no-hold measurements were larger, but even the combination of four no-hold tests explained only 35% of the variance of the combined hold measure. Overall, the level of unexplained variance in performance across hold and no-hold tests in our neurologically healthy sample cautions against the viability of using this method for accurately predicting premorbid ability in cognitively impaired patients.

Estimates based on word reading (NART and WTAR)

Significantly better performance was observed on the WTAR than the NART [$t(91) = 19.98, p < .001$], indicating both that the NART is the more difficult test, and that discrimination among more cognitively capable individuals on the basis of WTAR

Table 2. Correlations and direct comparison among hold and no hold measures.

Correlations	Hold tests			Pairwise <i>t</i> -tests
	Vocabulary	Information	Combined	
No-hold tests				
<i>Block Design</i>	.29	.32	.34	Hold vs. no-hold (two tests): <i>t</i> (91) = 2.16, <i>p</i> = .034
<i>Digit Span</i>	.40	.23	.36	
<i>Combined (2 tests)</i>	.49	.38	.49	
<i>Arithmetic</i>	.43	.32	.42	Hold vs. no-hold (four tests): <i>t</i> (91) = 2.34, <i>p</i> = .021
<i>Coding</i>	.39	.34	.41	
<i>Combined (4 tests)</i>	.58	.46	.59	

Note: *p* values not corrected for multiple comparisons.

performance may be problematic as a result of possible ceiling effects (Table 3). Performance across the WAIS-IV measures also differed significantly [$F(3, 272.59^1) = 3.12, p = .026$], although pairwise comparisons revealed that only one effect remained significant following Bonferroni correction, with FSIQ higher than PSI ($p = .043$).

NART and WTAR raw error scores exhibited a large correlation [$r(90) = .88, p < .001$] and both measures also showed significant negative correlations with age [$r(90) = -.64$ and $-.54, p < .001$, for NART and WTAR respectively]. Table 4 provides correlations of these test scores with WAIS-IV FSIQ, constituent indices and core subtest scaled scores. The main NART/WAIS-IV correlations and regression equations have previously been published (Bright et al., 2016) but have been included to facilitate comparison with WTAR and alternative methods presented here. Statistically, the tests provided equally precise predictions of WAIS-IV performance, with the strongest effects observed for FSIQ, GAI and VCI. Weaker correlations were observed against WMI and PRI. Correlations with PSI were comparatively poor, indicating that estimation of basic information processing speed should not be inferred on the basis of NART or WTAR scores. We also assessed the correlation between the mini-NART (McGrory et al., 2015) and WAIS-IV FSIQ, which had the effect of significantly reducing the correlation from $r(90) = .69$ to $r(90) = .63$ ($z = 2.41, p = .01$).

The range of NART-derived FSIQ predicted values in our sample was 43 IQ points, with our regression analysis revealing that the full distribution of possible predicted values ranged from 78 (50 NART errors) to 126 (0 NART errors). Point-by-point comparison against predicted WAIS and WAIS-R IQs included in the British NART-R test manual shows similar estimates at the high end of the distribution (but lowest for WAIS-IV), with estimates at the lower end falling between the WAIS (higher) and WAIS-R (lower) FSIQ estimates (Figure 2). The sample range was lower in our WTAR data, with 33 predicted FSIQ values, but the regression analysis revealed a wider distribution of estimates ranging from 59 (50 WTAR errors) to 120 (0 WTAR errors). Nevertheless, the scarcity of very low WTAR scores in our sample suggests that these lower FSIQ estimates should be interpreted with caution. The regression equations were as follows:

1. NART predicted WAIS-IV FSIQ = $-.9775 \times \text{NART error} + 126.41$
2. WTAR predicted WAIS-IV FSIQ = $-1.2206 \times \text{WTAR error} + 119.63$

Table 3. NART and WTAR raw error and predicted and observed WAIS-IV performance.

	Range	Mean	SD	Sig. diff.
NART errors (max = 50)	2–46	18.30	8.98	$p < .001$
NART predicted FSIQ	81–124	108.52	8.78	
WTAR errors (max = 50)	1–34	9.10	6.97	
WTAR predicted FSIQ	78–120	108.53	8.51	
Actual WAIS-IV performance				
FSIQ	80–150	108.52	12.71	$p = .043$
GAI	80–148	108.83	13.50	
VCI	74–138	107.14	14.73	
PRI	82–144	108.75	12.48	
WMI	80–148	106.07	14.28	
PSI	81–146	104.86	13.25	

NART, National Adult Reading Test; WTAR, Wechsler Test of Adult Reading; WAIS-IV Wechsler Adult Intelligence Scale – Fourth Edition; FSIQ, WAIS-IV full-scale IQ; GAI, General Ability Index; VCI, Verbal Comprehension; PRI, Perceptual Reasoning; WMI, Working Memory; PSI, Processing speed.

Table 4. Correlations of NART and WTAR performance with WAIS-IV FSIQ, index and subtest scores.

Measure	NART	WTAR	NART + WTAR
Full Scale IQ	.69***	.67***	.70***
Global Ability Index	.64***	.62***	.65***
Verbal Comprehension Index	.66***	.68***	.69***
Similarities	.36***	.44***	.41***
Vocabulary	.75***	.75***	.78***
Information	.53***	.57***	.56***
Perceptual Reasoning Index	.45***	.39***	.44***
Block Design	.29**	.23*	.27**
Matrix Reasoning	.43***	.38***	.42***
Visual Puzzles	.35**	.31**	.34***
Working Memory Index	.50***	.47***	.50***
Digit Span	.45***	.41***	.46***
Arithmetic	.40***	.39***	.41***
Processing speed Index	.36***	.36***	.37***
Symbol Search	.28**	.30**	.30**
Coding	.39***	.37***	.39***

NART, National Adult Reading Test; WTAR, Wechsler Test of Adult Reading; WAIS-IV Wechsler Adult Intelligence Scale – Fourth Edition; FSIQ, WAIS-IV full-scale IQ; *** $p < .001$; ** $p < .01$.

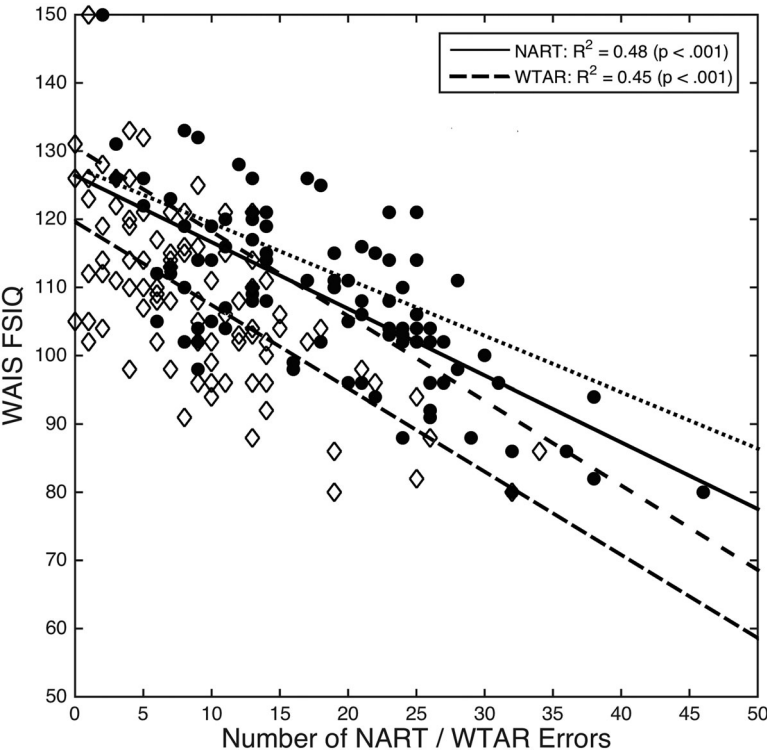


Figure 2. Linear correlation between National Adult Reading Test/Wechsler Test of Adult Reading (NART/WTAR) errors and Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV) full-scale IQ (FSIQ). The original published estimates of WAIS (dotted) and WAIS-R FSIQ (wide-space dashed) from the manual (Nelson & Willison, 1991) are included for comparison.

We computed regression equations for NART and WTAR scores against each of the WAIS-IV indices (excluding PSI, which was poorly correlated, as described above). Figure 3 presents scatterplots relating NART error to index scores. NART consistently produced higher WAIS-IV estimates than WTAR for a given level of performance, with the level of disparity increasing as a function of error. The regression equations were as follows:

NART:

Predicted General Ability Index (GAI) = $-.9656 \times \text{NART errors} + 126.5$

Predicted Verbal Comprehension Index (VCI) = $-1.0745 \times \text{NART errors} + 126.81$

Perceptual Reasoning Index (PRI) = $-.6242 \times \text{NART errors} + 120.18$

Working Memory Index (WMI) = $-.7901 \times \text{NART errors} + 120.53$

WTAR:

Predicted General Ability Index (GAI) = $-1.2025 \times \text{WTAR errors} + 119.77$

Predicted Verbal Comprehension Index (VCI) = $-1.4411 \times \text{WTAR errors} + 120.25$

Perceptual Reasoning Index (PRI) = $-.6931 \times \text{WTAR errors} + 115.06$

Working Memory Index (WMI) = $-.9579 \times \text{WTAR errors} + 114.78$

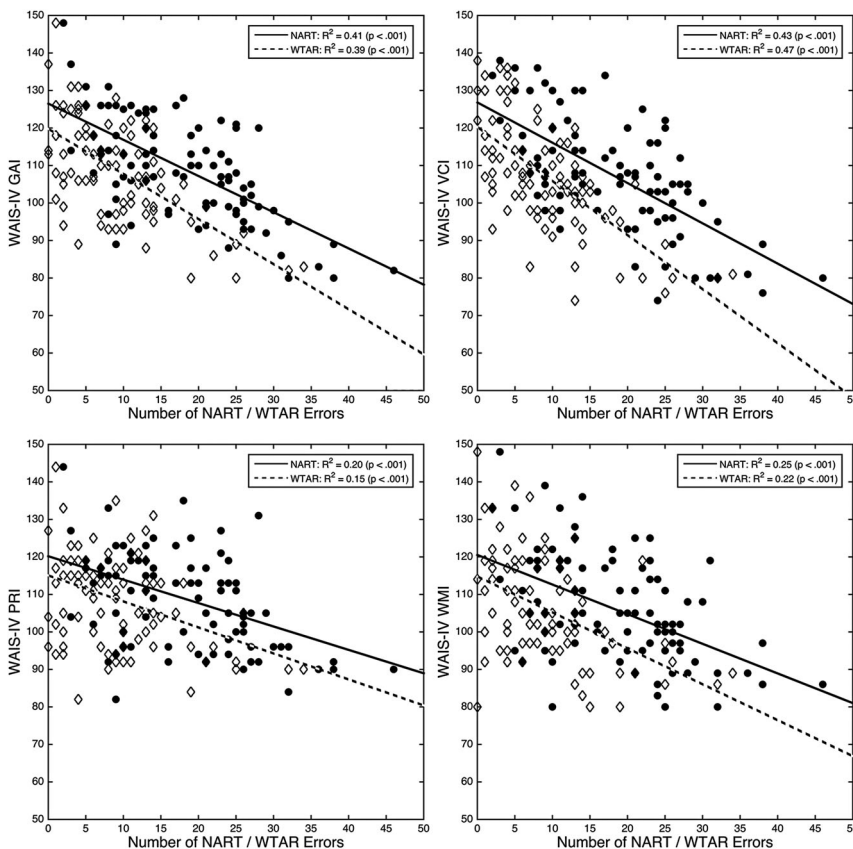


Figure 3. Scatterplots showing linear correlations relating number of the National Adult Reading Test (NART) and Wechsler Test of Adult Reading (WTAR) errors to (A) General Ability Index (GAI); (B) Verbal Comprehension (VCI); (C) Perceptual Reasoning (PRI); and (D) Working Memory (WMI). Processing speed (PSI) has been excluded.

Estimates based on combined test and demographic data

Linear regression models were used to determine the effect of combining test and demographic data on the accuracy of our estimates of WAIS-IV performance. Stepwise regression using standard inclusion ($p = .05$) and exclusion ($p = .1$) criteria indicated that the best model in all cases contained two predictor variables (with the demographic variable explaining an additional 5% of the variance in FSIQ scores). This was the case for equations incorporating NART, WTAR, and the sum of these test scores (Table 5). The benefit of including the sum of NART and WTAR errors on estimation accuracy was negligible. Age significantly improved the precision of FSIQ estimates based on NART and total NART + WTAR performance, and education improved WTAR-derived estimates only. The two variable equations are as follows:

NART: estimated FSIQ = $141.126 - (1.26 \times \text{NART error}) - (.236 \times \text{age})$

WTAR: estimated FSIQ = $111.553 - (1.087 \times \text{WTAR error}) + (2.976 \times \text{education})$

NART + WTAR: estimated FSIQ = $136.839 - (.720 \times (\text{NART} + \text{WTAR error})) - (.212 \times \text{age})$

Table 6 provides FSIQ estimates on the basis of the single and two variable models at three levels of the relevant demographic measure. Inclusion of age with NART provided an additional potential benefit beyond the improved precision of estimate, by extending the range of possible FSIQ values at both ends of the distribution. Inclusion of education with WTAR is more problematic, since we cannot know what the maximum educational level achieved will be for the younger participants in our sample (i.e., some participants were in full-time education and/or may not have reached their peak level of achievement at the time of testing).

Table 5. Linear regression models incorporating test scores (NART, WTAR) and demographic variables as predictors of WAIS-IV FSIQ performance.

Model	NART	WTAR	Age	Education	R^2
NART, demographics					
1	1	0	0	0	.48
2	1	0	1	0	.53
3	1	0	0	1	.51
4	1	0	1	1	.54
WTAR, demographics					
5	0	1	0	0	.45
6	0	1	1	0	.46
7	0	1	0	1	.50
8	0	1	1	1	.50
NART, WTAR, demographics					
9	1	1	0	0	.49
10	1	1	1	0	.54
11	1	1	0	1	.55
12	1	1	1	1	.55
Demographics only					
13	0	0	1	0	.07
14	0	0	0	1	.18
15	0	0	1	1	.27

NART, National Adult Reading Test; WTAR, Wechsler Test of Adult Reading; WAIS-IV Wechsler Adult Intelligence Scale – Fourth Edition. FSIQ, WAIS-IV full-scale IQ; Note: 1 = included in model; 0 = excluded from model. Bold values indicate significant single predictor models and stepwise multivariate models in which the fit is significantly improved.

Table 6. Single test (model 1) and combined (model 2) example estimates of WAIS-IV FSIQ.

		Model 2			
		Age (years)			
NART errors	Model 1	20	45	70	
0	126	135	131		123
25	102	105	99		93
50	78	73	68		62
WTAR errors		Model 2			
		Education level			
		4	3	2	1
0	120	123	120	118	115
25	89	96	93	90	87
50	59	69	66	63	60

NART, National Adult Reading Test; WTAR, Wechsler Test of Adult Reading; WAIS-IV, Wechsler Adult Intelligence Scale – Fourth Edition. FSIQ, WAIS-IV full-scale IQ; Note: Education level 1 = GCSE/equivalent or below; 2 = A level/equivalent; 3 undergraduate degree; 4 postgraduate degree.

Discussion

Clinicians and researchers have at their disposal a range of methods for the estimation of premorbid cognitive ability, and their choice of method will be informed by the characteristics of the presenting patient and their own expertise and experience. Each method has strengths and weaknesses. For example, performance on tests such as the NART and WTAR is unlikely to be entirely insensitive to neurological impairment, and the degree of sensitivity is likely to differ from one patient and/or condition to another. Such tests also require neuropsychological assessment skills/training, take time to administer, and can contribute to patient fatigue. These potential problems can be avoided by eschewing estimates based on current test performance, i.e., by using demographic data only, but demographic-based approaches raise other concerns. Categories based on occupational status and education, for example, are arguably too coarse to provide an accurate premorbid IQ for a specific individual. Best performance and embedded hold/no-hold methods are also problematic. Wide variability is observed in performance across subtests in intelligence batteries, along with poor inter-test correlations.

Despite the considerable limitations associated with all currently available methods, even the most experienced clinician would be constraining his or her ability to deliver optimal clinical management of a presenting neurological patient if estimation of premorbid ability was not attempted. In practice, the clinician considers evidence from multiple sources when estimating the degree of cognitive impairment (if any), but to avoid bias and constrain subjectivity, it is crucial to employ evidence-based assessment approaches in this process (e.g., Youngstrom, Choukas-Bradley, Calhoun, & Jensen-Doss, 2015). Our findings suggest that tests of word reading/vocabulary knowledge provide the most reliable and precise estimates of WAIS-IV performance, and previous work indicates that their utility for predicting premorbid IQ holds in a range of neurological conditions (Bright et al., 2002). However, we also found that predictive accuracy can be modestly but significantly improved through the use of combined test scores with demographic information (NART with age, and WTAR with education). Since the NART (and NART-R) were published, similar tests of reading/vocabulary knowledge have also been proposed that provide predicted scores incorporating one or more

demographic variables (the WTAR against WAIS-III and the TOPF against WAIS-IV). The value of the NART and WTAR for estimating WAIS-IV index scores is more questionable, showing large correlations with the VCI and GAI but relatively modest correlations with WMI and PRI, suggesting that caution should be employed in drawing inferences about premorbid executive function and fluid ability. Consistent with these findings were the large correlations between test performance and age, indicating that both the NART and WTAR tap “crystallised” knowledge (which typically improves across our sample age range) rather than fluid ability (which typically peaks in early adulthood and subsequently declines; Cattell, 1971). These tests should not be used to infer premorbid processing speed.

The published NART/NART-R manual provides estimates of WAIS or WAIS-R performance, and the WTAR presents WAIS-III estimates, all of which are now obsolete. Researchers and clinicians working with UK populations who employ NART or WTAR may therefore wish to consider applying our equations in order to compare actual and predicted premorbid WAIS-IV (rather than WAIS-R/WAIS-III) performance. Approaches based on the NART, in particular, remain popular with many researchers and clinicians in the UK, USA, Canada and Australia, but even though the Test of Premorbid Function (TOPF) was designed to supersede the WTAR, the WTAR remains widely used. Field work is currently underway to develop WAIS-V, which, once published, will require the development of new standardised estimates if use of the NART or WTAR is to continue.

Directions for future research

The development of standardised tools such as the NART and WTAR has undoubtedly improved the ability to predict meaningful baseline levels of performance so that the impact of a neurological condition on cognition can be judged. Nevertheless, we question the ambition of the tools developed to date and encourage the development of novel approaches to improving premorbid estimates. For example, both the NART and the WTAR use equal weightings for each of the 50-test items comprising each test. With large samples, however, reliable stimulus-specific coefficients can be computed in which the predictive value of each stimulus is individually weighted. Such scaling techniques may provide the basis for dramatic and highly significant increases in predictive power – in our data, for example, we observed a 46% increase in the variance shared between rescaled NART values and WAIS-IV FSIQ. They may also identify redundant test items that possess little, if any, predictive power. However, such methods typically require large datasets and replication studies – and for this reason we have not presented these statistics here.

The extent to which specific disorders may impact on those abilities assessed with tests such as the NART or WTAR is difficult to predict, particularly for more severely impaired patients or those with language and/or semantic memory impairment, and more work is required in this area. Development of methods for estimation of premorbid functioning in cognitive domains other than IQ may also be beneficial in supporting clinical judgement by providing more direct comparison against presenting symptoms (whether memory loss, deterioration in conceptual knowledge, executive dysfunction, or other reported deficits). In the present study, for example, NART and WTAR performance was only moderately sensitive to current working memory and perceptual reasoning ability, implying limited utility of such tests for estimating premorbid nonverbal/

fluid intelligence in neurological patients. By definition, psychometric intelligence predicts performance across all cognitive domains, but in practice such generalised inferences are likely to be problematic in many cases. Future studies should aim to identify methods optimally adapted to specific conditions, so that, to the greatest extent possible, like is compared with like.

Endnotes

1. Degrees of freedom corrected for violation of sphericity assumption using the Greenhouse-Geisser method.

Acknowledgements

We wish to thank Emily Hale, Vikki Jane Gooch and Thomas Myhill for their help with data collection.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

References

- Barona, A., Reynolds, C. R., & Chastain, R. (1984). A demographically based index of premorbid intelligence for the WAIS—R. *Journal of Consulting and Clinical Psychology*, 52(5), 885–887. <https://doi.org/10.1037/0022-006X.52.5.885>
- Bright, P., Hale, E., Gooch, V. J., Myhill, T. D., & van der Linde, I. (2016). The National Adult Reading Test: Restandardisation against the Wechsler Adult Intelligence Scale – Fourth Edition. *Neuropsychological Rehabilitation*. <http://doi.org/10.1080/09602011.2016.1231121>
- Bright, P., Jaldow, E., & Kopelman, M. D. (2002). The national adult reading test as a measure of premorbid intelligence: A comparison with estimates derived from demographic variables. *Journal of the International Neuropsychological Society*, 8, 847–854. <https://doi.org/10.1017/S1355617702860131>
- Cattell, R. B. (1971). *Abilities: Their structure, growth and action*. Boston: Houghton-Mifflin.
- Crawford, J. R., & Allan, K. M. (1997). Estimating premorbid WAIS–R IQ with demographic variables: Regression equations derived from a UK sample. *Clinical Neuropsychologist*, 11, 192–197. <https://doi.org/10.1080/13854049708407050>
- Crawford, J. R., Deary, I., Starr, J., & Whalley, L. (2001). The NART as an index of prior intellectual functioning: A retrospective validity study covering a 66-year interval. *Psychological Medicine*, 31, 451–458. <https://doi.org/10.1017/S0033291701003634>
- Crawford, J. R., Nelson, H. E., Blackmore, L., Cochrane, R. H. B., & Allan, K. M. (1990). Estimating premorbid intelligence by combining the NART and demographic variables: An examination of the NART standardisation sample and supplementary equations. *Personality and Individual Differences*, 11, 1153–1157. [https://doi.org/10.1016/0191-8869\(90\)90028-P](https://doi.org/10.1016/0191-8869(90)90028-P)
- Crawford, J. R., Stewart, L. E., Cochrane, R. H. B., Parker, D. M., & Besson, J. A. O. (1989). Construct validity of the national adult reading test: A factor analytic study. *Personality and Individual Differences*, 10(5), 585–587. [https://doi.org/10.1016/0191-8869\(89\)90043-3](https://doi.org/10.1016/0191-8869(89)90043-3)
- Donders, J., Tulskey, D. S., & Zhu, J. (2001). Criterion validity of new WAIS–III subtest scores after traumatic brain injury. *Journal of the International Neuropsychological Society*, 7(07), 892–898.
- Franzen, M. D., Burgess, E. J., & Smith-Seemiller, L. (1997). Methods of estimating premorbid functioning. *Archives of Clinical Neuropsychology*, 12(8), 711–738.
- Griffin, S. L., Mindt, M. R., Rankin, E. J., Ritchie, A. J., & Scott, J. G. (2002). Estimating premorbid intelligence: Comparison of traditional and contemporary methods across the intelligence continuum. *Archives of Clinical Neuropsychology*, 17(5), 497–507. [https://doi.org/10.1016/S0887-6177\(01\)00136-6](https://doi.org/10.1016/S0887-6177(01)00136-6)
- Groth-Marnat, G., & Wright, A. J. (2016). *Handbook of psychological assessment* (6th ed.). Hoboken, NJ: John Wiley & Sons.
- Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York, NY: Oxford University Press.

- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). New York, NY: Oxford University Press.
- Mathias, J. L., Bowden, S. C., & Barrett-Woodbridge, M. (2007). Accuracy of the Wechsler Test of Adult Reading (WTAR) and National Adult Reading Test (NART) when estimating IQ in a healthy Australian sample. *Australian Psychologist*, 42(1), 49–56. <https://doi.org/10.1080/00050060600827599>
- McGrory, S., Austin, E. J., Shenkin, S. D., Starr, J. M., & Deary, I. J. (2015). From “aisle” to “labile”: A hierarchical National Adult Reading Test scale revealed by Mokken scaling. *Psychological Assessment*, 27(3), 932–943. <https://doi.org/10.1037/pas0000091>
- Mortensen, E. L., Gade, A., & Reinisch, J. M. (1991). A critical note on Lezak’s “best performance method” in clinical neuropsychology. *Journal of Clinical and Experimental Psychology*, 13, 361–371. <http://doi.org/10.1080/01688639108401050>
- Nelson, H. E. (1982). *National Adult Reading Test (NART): For the assessment of premorbid intelligence in patients with dementia: Test manual*. Windsor: NFER-Nelson.
- Nelson, H. E., & O’Connell, A. (1978). Dementia: The estimation of premorbid intelligence levels using the New Adult Reading Test. *Cortex*, 14(2), 234–244.
- Nelson, H. E., & Willison, J. (1991). *National Adult Reading Test (NART)*. Windsor: NFER-Nelson.
- Office of Population, Censuses and Surveys. (1980). *Classification of occupations*. London: HMSO.
- Pearson, N. C. S. (2009). *Advanced clinical solutions for WAIS-IV and WMS-IV: Administration and scoring manual*. San Antonio: The Psychological Corporation.
- Powell, B. D., Brossart, D. F., & Reynolds, C. R. (2003). Evaluation of the accuracy of two regression-based methods for estimating premorbid IQ. *Archives of Clinical Neuropsychology*, 18(3), 277–292. [https://doi.org/10.1016/S0887-6177\(02\)00135-X](https://doi.org/10.1016/S0887-6177(02)00135-X)
- Reynolds, C. R. (1997). Postscripts on premorbid ability estimation: Conceptual addenda and a few words on alternative and conditional approaches. *Archives of Clinical Neuropsychology*, 12(8), 769–778. [https://doi.org/10.1016/S0887-6177\(97\)00051-6](https://doi.org/10.1016/S0887-6177(97)00051-6)
- Scott, J. G., Krull, K. R., Williamson, D. J., Adams, R. L., & Iverson, G. L. (1997). Oklahoma premorbid intelligence estimation (OPIE): Utilization in clinical samples. *The Clinical Neuropsychologist*, 11(2), 146–154. <https://doi.org/10.1080/13854049708407043>
- Skilbeck, C., Dean, T., Thomas, M., & Slatyer, M. (2013). Impaired National Adult Reading Test (NART) performance in traumatic brain injury. *Neuropsychological Rehabilitation*, 23(2), 234–255. <https://doi.org/10.1080/09602011.2012.747968>
- Vanderploeg, R. D., Schinka, J. A., & Axelrod, B. N. (1996). Estimation of WAIS—R premorbid intelligence: Current ability and demographic data used in a best-performance fashion. *Psychological Assessment*, 8(4), 404–411. <https://doi.org/10.1037/1040-3590.8.4.404>
- Wechsler, D. (1958). *The measurement and appraisal of adult intelligence*. Baltimore: Williams & Wilkins.
- Wechsler, D. (1997). *WAIS-III: Wechsler Adult Intelligence Scale*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2001). *Wechsler Test of Adult Reading: WTAR*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2008). *Wechsler Adult Intelligence Scale* (4th ed.). San Antonio, TX: Pearson Assessment.
- Wechsler, D. (2011). *Test of Premorbid Functioning. UK version (TOPF UK)*. London: Pearson Assessment.
- Youngstrom, E. A., Choukas-Bradley, S., Calhoun, C. D., & Jensen-Doss, A. (2015). Clinical guide to the evidence-based assessment approach to diagnosis and treatment. *Cognitive and Behavioral Practice*, 22(1), 20–35. <https://doi.org/10.1016/j.cbpra.2013.12.005>